CANTWELL

500-81-0053

THE URBAN INSTITUTE

2100 M Street, N.W.
Washington, D.C. 20037

Project Report

REPORTS
R
728.
.5
P74
1981

R
728.5
.P74
1981

3075-1 November 1981

Preliminary Research Plan for Alternative Methods for Developing a Relative Value Scale of Physicians' Fees

Jack Hadley, Philip Held,
John Holahan, and Judith Wagner

The research for this report was supported by Contract No. HCFA-500-81-0053 from ORDS/HCFA/DHHS. Opinions expressed are those of the authors and do not necessarily represent the views of the Urban Institute or its sponsors.

THE URBAN INSTITUTE
2100 M Street, N.W., Washington, D. C. 20037

I. INTRODUCTION

This report provides a systematic, though preliminary review of alternative methods of constructing relative value scales. The report's primary goal is to facilitate an initial comparison of methods in order to determine which methods merit more detailed review and analysis. Six general approaches, listed in Figure 1, are covered. (The total number of variants within these general approaches is potentially much larger.) Separate sections follow a similar format to report information about each method. The concluding section makes recommendations regarding which methods are most promising within the constraints of this study.

The first method, the relative charge approach, is based on data on physicians' charges or fees. Physicians' charges for each procedure are arrayed into distributions and a point on each distribution (mean, mode, median, or any other percentile) is chosen to represent that procedure's value. Relative values are constructed by computing the ratio of each procedure's value to some numeraire.

The next three approaches are members of the cost-based family. The least complicated focuses on physicians' time as the basis for constructing relative values. The amount of time the physician spends in performing each procedure, plus possibly some base value measure to account for the costs of other inputs, represents each procedure's value. Micro-costing, which includes time/motion and task analysis, attempts to make a more complete accounting of the cost of performing each procedure. The contributions of all inputs and their costs are measured and added up to compute the full resource cost of performing each procedure. Relative values are then constructed by computing ratios of each procedure's full cost to the cost of a numeraire

Figure 1

General Methods of Constructing Relative Value Scales

<u>Charge (Fee) Methods</u>

Relative Charges

<u>Cost Methods</u>

Physicians' Time (plus Base Value)

Micro-Costing (Time/Motion, Work Sampling)

Statistical Cost Functions

<u>Consensus Methods</u>

Delphi

Social Preferences

procedure. Finally, statistical cost functions rely on the principles of economic theory and multivariate statistical techniques in order to estimate what each procedure's value should be under certain ideal conditions.

The last two approaches we will examine are different consensus techniques. One is the delphi method, which is an iterative, panel-of-experts approach. In essence, the panel assigns relative values to procedures based on individual members' knowledge and preferences, the information presented to the panel, and the panel's internal procedures. The social preference approach also assigns (or adjusts) relative values based on some type of consensus (as opposed to a specific statistical method). It differs from the delphi method in that the consensus (statement of preferences) need not be arrived at through a formal panel process. Rather, the identification of social preferences can be based on a much braoder set of inputs, e.g., Congressional and other legislative documents or statements, executive department positions, and provider and patient groups' representatives. Two good examples of the social preference method are the target expenditure and target rate of return approaches used in West Germany and Quebec, Canada, respectively. California's manipulation of the structure of Medicaid fees is another example.

The primary goal of the initial phase of the project is to describe and compare the alternative methods along two major dimensions. One is consistency with several broad objectives of a relative value scale. In particular, an RVS should convey proper signals to both physicians and consumers about resource allocation; it should be usable, in conjunction with a procedure coding terminology, by insurers to pay claims; it should provide a basis for setting fees or maximum allowable payments; it should discourage artificial manipulation or "gaming" by providers; it should provide the potential for auditing and for monitoring providers; and it should enhance, or at least not impede, competition among providers. The extent to which alternative methods meet these objectives will be an important criterion for comparing approaches.

The second major dimension covers a variety of technical issues pertinent to actually constructing an RVS. For example, what data are needed? How readily available are they? How costly? Are sophisticated analytic methods required? How feasible is local implementation of the methods by individual carriers and intermediaries? How difficult is it to modify or update the RVS? How well does it accommodate new procedures, infrequently performed procedures, and procedures tied to rapidly changing technologies? How costly would it be to construct specialty-specific, community-size-specific, and/or procedure-type-specific RVSs?

In a later stage of the project we will examine the cost implications and the distributional implication of alternative methods and of alternative decisions regarding some of the technical issues. The examination of the cost implications will attempt to consider both program (Medicare and Medicaid) costs and total costs. The analysis of distributional implications will focus on payments to physicians in different specialties and locations, and payments on behalf of different classes of beneficiaries (by age, sex, and possibly class of eligibility). These simulations will provide a better estimate of the differences among alternative approaches and the consequences of different technical decisions one could make in implementing any particular approach.

In conducting this study it is important to remember that no single approach is likely to dominate all others along all possible evaluation criteria. The "best" approach will probably be different for different types of procedures and different situations. Consequently, the optimal method of constructing an RVS may require both concatenating methods and using them jointly. This suggests that future stages of the project will probably concentrate on applying certain methods to only some classes of procedures.

The preliminary discussions which comprise the next section follow a similar format. Each begins with a brief description of the method's basic elements and information on whether the method has ever been used to actually construct an RVS. Data and technical analysis requirements are discussed next. What kind of data are needed? What kind of data are available? How costly would it be to obtain the needed data? What kinds of people need to do what kinds of things to convert the data into an RVS?

We then examine some of the RVS's characteristics. Does it reflect physicians' costs, social costs, physicians' charges, patients' valuations, society's valuations, or some mix of these? Can it be updated easily? What are its problem procedures? Can individual carriers implement it easily? Can area-, specialty-, or community-size-specific RVS's be constructed? What procedures is it best suited for? Finally, a summary of the method's strengths and weaknesses is presented.

The concluding section of this report presents our recommendations regarding further analysis of each method.

II. PRELIMINARY EXAMINATION OF ALTERNATIVE METHODS

RELATIVE CHARGES

Description

The best-known set of relative value scales is probably the one developed by the California Medical Association between the 1950s and 1974. Although specific methodologies varied over time, the essence of the California Relative Value Study (CRVS) approach was to base relative values on physicians' charges. The initial CRVS conducted a survey of what physicians generally charge. The most recent version was based on claims submitted to third-party payers. Thus, there is ample historical precedent for basing relative values on charges.

The essence of the charge-based approach is that physicians' charges for each procedure are arrayed into a distribution and some point on each distribution is chosen to represent the procedure's "value". Relative values are constructed by computing the ratio of each procedure's value to that of a numeraire procedure. Although conceptually straightforward, the charge-based approach has numerous variations, all of which may lead to different RVSs. What point on the distribution should be chosen? Which physicians' charges should be combined? How should they be combined to form a distribution?

Data Requirements

In order to explore the consequences and implications of these types of choices, it is important to have a large and flexible data base. Charge data are potentially available from private insurers, Medicare carriers, and Medicaid programs in every state in the nation. The data could be used to reveal how physicians in various specialties and various localities price each procedure relative to others.

Three data bases are currently available to us for developing an RVS with charge data. Each has considerable advantages but some disadvantages. The first is the Urban Institute California Medicare and Medicaid claims data file. This data base, which has been used in previous research done for ORDS/HCFA, contains all claims submitted to Medicare and Medicaid by approximately 7,500 California physicians during the first calendar quarter of each year between 1973 and 1978. We would most likely use only 1978 data. The data base contains charges for all procedures (medical, surgical, laboratory, and radiology) billed by eleven major specialties including new and old physicians, group and solo physicians, urban and rural physicians, and foreign and domestically trained physicians. The data set would allow us considerable flexibility in choosing many different percentiles on the distribution of charges and in examining differences in relative values among specialties and geographic areas.

The second data base is the charge data provided by each Medicare carrier to establish the 1982 prevailing charge screens. Carriers were required to provide the 50th percentile and the unindexed 75th percentiles for each of the 100 procedures for which Medicare prevailing charge data are natinally collected. These data offer the advantage of covering a large number of procedures provided by many specialties. More importantly, it would permit us to test for differences in relative values in different parts of the country.

The third data file is the Health Insurance Association of America's survey of twenty-five commercial insurance companies' paid claims for surgical procedures. The file is produced every six months using data for the previous twelve-month period. Procedures are identified by their CRVS code. For each of approximately 250 zip codes, the file reports the number of charges received, the mean and model charges, and various percentiles between 50 and

95. Although this file covers only surgical procedures, it has the advantage of not being limited to Medicare and Medicaid claims. (Many Medicare prevailings may, of course, have been computed from data which included carriers' private business.) Thus, it provides a good opportunity to compare surgical RVSs built from private data with those based on public or mixed public/private data.

In general, the data for developing an RVS based on charges are available through claims submitted for physicians. Processing costs are high. Charges for each procedure must be arrayed, different percentiles found, and then relativity determined. These steps require multiple passes of large volumes of physicians' claims. Datas acquisition costs, however, are relatively low because physicians must submit claims to be reimbursed for their services. Furthermore, most carriers routinely computerize claims in order to process payments.

Technical Analysis Requirements

Implementing a charge-based reimbursement system requires substantial computing capacity and sophisticated data processing support. Personnel with advanced training in statistics, economics, or health care financing are not likely to be needed.

Characteristics of a Charge-Based RVS

The resulting RVS is likely to reflect costs of providing each procedure or each service within procedure types. It may not very well reflect costs across procedure types or patients' valuations. This is true because of potential biases introduced by the different levels of insurance coverage people have for different kinds of procedures. Historically, office visits have not been well insured. Surgical procedures and diagnostic services, such as laboratory and x-ray services, have been well insured. As a result, office visit fees have not increased at the same rates as surgery, laboratory, and radiology fees. Thus a charge-based RVS system is likely to overstate the relative costs as well as the relative values of the surgery, laboratory and x-ray procedures and undervalue office visit procedures. This is also true of any other procedures performed that have not been well insured in the past.

Updating

While data currently exist to develop RVSs on the basis of physicians' charges, it may be more difficult to implement and update a charge-based RVS if Medicare were to adopt a fee schedule. If the fee schedule were constructed from an RVS based on charges, it could make use of claims (charge) data in the first year. However subsequent charge data would not be as useful because they most likely would be affected by the allowed fees. Medicare would have to make use of private insurance data as input in updating or altering its RVS in subsequent years, or to rely on methods other than charges for periodic modifications of relative values.

Pricing Problem Procedures

There are no particular advantages to a charge-based system in pricing problem procedures. Charges for new procedures may overstate their true marginal costs if the costs of research and development are loaded into the charges for the initial period of performance. They are also unlikely to be a good measure of patients' valuation because of the lack of experience with the procedure as well as, in many cases, insurers' automatic acceptance of the charges set by providers.

Charge-based systems are likely to do a fairly poor job of pricing procedures where the technology is undergoing rapid change. For example, if technological innovations reduce the price of performing certain procedures, it is likely that charges will not fall in response to the lower costs. Because of insurance, prices are likely to remain high even if costs fall.

Specialty-and Locality-Specific RVSs

Charges represent perhaps an ideal system for developing different RVSs for different specialties and for different regions or geographic areas. Because charges are available for each physician in each specialty in each region, it provides a mammoth data base in which alternative specialty-specific or locality-specific RVSs can be developed.

Implementation by Individual Carriers

Individual carriers could implement a charge-based method relatively easily. It is unlikely, however, that Medicare would want the carriers to bear this cost because of the large amount of duplication that would occur.

Suitability Across Procedures

A charge-based RVS system seems suitable for all classes of procedures. As noted above, however, it provides a better guide to relative costs, values, etc. within procedure types, that is within medicine, surgery, radiology and laboratory. It does not provide a good guide for relative values or costs across procedure types for reasons stated above.

External Costs

Charge-based systems do not take account of costs outside the physician's practice, either the costs to patients, or costs borne by the hospital. For example, the relative value of a cesearean section delivery is likely to be greater than that of a vaginal delivery, since physicians' charges are typically greater for the former than the latter. An insurer may wish to "devalue" cesarean deliveries relative to vaginal deliveries, however, because cesareans require long hospital stays and more restricted activity days for the mother than do vaginal deliveries. (The higher relative value of cesarean sections may also reflect insurance distortions, since it is almost always covered as a hospital procedure while vaginal deliveries are generally less well covered.) These concerns are relevant if HCFA is concerned about total costs rather than only payments to physicians.

Strengths

The major strength of a charge-based RVS system is the large amount of data that and potentially available for developing RVSs which can vary by locality and by specialty. The data permit an observation of the real world, that is the way large numbers of physicians price individual procedures. Aberrations resulting from an individual physician's pricing behavior will not affect the development of a relative value scale because data are available for large numbers of physicians.

Weaknesses

Its primary weakness is the bias introduced because charges have been influenced by the availability of insurance. Well-insured procedures are likely to be overpriced; the poorly insured procedures to be underpriced. The charge-based system would also be costly to update and alter over time because it will require continual data processing and analysis efforts on large scale data bases to determine if relative prices have changed.

PHYSICIANS' TIME PLUS BASE VALUE

Description

As implied by the name, this method of RVS construction is based on the amount of physician time required for a procedure. Estimates of physician time can be obtained by direct observation, by work activity logs maintained on a prospective basis by physicians, or by retrospective surveys of physicians' practices. Though each of these data sources has potential biases, any one of them allows the estimation of the mean as well as the distribution of procedure time.

Although it is not always clear how to define the endpoints i.e., the beginning and end of a procedure, the use of relative physician time provides a precise basis for measurement. The measure is easy to understand and is therefore likely to be acceptable to physicians and insurors. Moreover, procedures that are most frequently performed in hospitals, such as inpatient surgery or nuclear medicine, may be more amenable to this approach, for outside of his or her own time, the physician bears little of the direct expense.

A fundamental issue regarding the use of such an approach to RVS construction is whether a resulting fee schedule would offer appropriate incentives to perform certain procedures. To the extent that procedure complexity is uncorrelated with procedure time, those that are unusually complex relative to their time requirements would be undervalued. Procedures requiring high investments in equipment or ancillary personnel would also be undervalued. The Mendenhall data (available through a subcontract with Lion Associates) could be used, however, to test the correlation of physicians' estimates of procedure complexity with procedure time.

A variation of the physicians' time approach is to build in a base value to which physicians' time is added. For example, the base might include the average fixed costs (overhead) of a physician's practices. (This is basically the approach used by Hsaio and Stason, 1979.) A more sophisticated method might include all non-physician costs in the base. Thus specialties which produce services that tend to employ relatively large numbers of ancillary personnel and relatively little physician time would not necessarily be undervalued. Other forms of this methodology emply modifiers in addition to the time plus base unit. For example, the anesthesia procedures of the CPT-4 are based on physician time and use a modifier which adjusts the RVS for patient complexity or severity.

An appealing approach to the use of physician time in constructing an RVS would be to consider diagnostic-specific RVSs, based on physician time and perhaps even including imputed values for ancillary services. For example, it would be possible to estimate the distribution of physician time reported for the treatment of the more common diagnostic conditions. This approach would have the advantage of focusing on visits, a procedure which accounts for a significant portion of total physician costs.

Data

There are basically two data sources which provide estimates of physician time that could be used to test and evaluate a methodology for RVSs based on physician time.

> Mendenhall - USC Data (available through Lion Associates) have physician time per patient encounter including diagnosis and severity. For example, the data are available for six specialties* and the sample sizes are quite large (approximately 40,000 patient encounters per specialty for the primary care physicians and

---

*General and family practice (two separate specialties), internal medicine, pediatrics, obstretrics and gynecology, and general surgery. The Mendenhall study did survey a greater number of specialties, but the "front end" costs of working with these data are very high. Of the six specialties we mention here, Lion Associates have five up and running and expect the sixth to be up by the end of the year.

general surgery. The data cover ambulatory and inpatient settings (including the time per surgical procedure. However, it is not known at this time whether preoperative, prep time is included.) The data, however, are limited in that they provide little or no information on the non-physician resources consumed. Finally, a drawback is that the survey was a one time crossection.

The National Ambulatory Medical Care Surveys (NAMCS) also have physician time per encounter including diagnosis and severity. The data also have the unique advantage of being part of an ongoing annual survey of reasonably large sample size.* Limitations of these data include restricted information on procedure classifications that are comparable to other RVS systems and a restriction to ambulatory settings only. At the least this restriction excludes most surgical procedures.

While the sample sizes of the NAMC survey are large, for some purposes, the samples will be fairly limiting. For example, specialty differences and visits by diagnosis will have marginal sample sizes at best.

The Mendenhall data are similar to NAMC data in that both have a diagnosis and both have a severity measure although the two measures are not identical. While the Mendenhall survey employed a complexity of service

*For example, in 1978, NAMCS had 47,291 patient physician encounters. Of this number, 36 percent were to a GP; 12 percent IM; 10 percent PED; 6 percent GS, 9 percent OBG and the rest distributed over 8 other specialties.

classification that is comparable to the CPT-4 office visit classification, there is nothing comparable in NAMC. Consequently, procedural classification with the NAMC data will be limited. The use of diagnostic classifications for visits is the principle correspondence between the NAMCS and Mendenhall data sources.

Data Costs and Requirements

Estimates of physician time are technically easy to obtain, but can of course be costly. There is, however, the ongoing NAMC survey which covers visits and the like in an adequate fashion. Data for inpatient procedures are not covered by NAMC and are consequently more of a problem and any new efforts in this area are likely to be fairly expensive.

A major problem with the existing data sources is the difficulty of obtaining the "base value," the estimates of severity/complexity or of the other resource requirements. It may be possible with the Mendenhall data to construct base values which reflect severity and/or complexity, since there are numerical ratings for these in the data. The NAMC survey could possibly be modified in future years to obtain some correlates of non-physician costs for office procedures such as non-physician staff and floor space. But these estimates would be average values for a practice and would not be able to capture the true differences in use of non-physician resources across procedures.

Technical Analysis Requirements

Given an appropriate data base and a developed methodology, middle level analysts could implement an RVS based on physician time.

Characteristics of an RVS Based on Physician Time

a. Does Physician Time Reflect Value?

A major issue of any RVS is whether the scale reflects values, fees, or costs or what. An RVS based on physician time will at best reflect costs and whether the relative costs across procedures are biased or not will depend on the particular procedure, physician specialty, and where the procedure was produced (in- or outpatient). RVSs for procedures which use no non-physician inputs or a quantity of inputs which are proportional to physician time will be likely to reflect relative costs reasonably well. For example, the RVS for anesthesia procedures provided in a hospital setting, where non-physician costs are excluded from physician reimbursement, may be a reasonable approximation of the relative costs. However, RVSs based on physician time for surgical procedures which can be performed in a physician's office or in an outpatient setting of a hospital may not accurately reflect physician costs since in the former the physician bears all costs while in the latter the physician may not be responsible for the hospital's costs.

Costs may not reflect value for a number of reasons. An example would be the differences in value which occur across geographic locations. Physician time will be more valuable in some places than in others and an RVS based on physician time may not reflect costs or values for similar services provided in different places. For example, the value of a surgical procedure for a condition which could be treated medically will be different in an area where there are few medical specialists.

b. Updating

Updating is possible but likely to be costly. If some questions can be "piggybacked" into the NAMC survey, then ambulatory RVSs could be updated at relatively low marginal costs. Updating inpatient surgical procedures would require developing a new periodic survey. However, data could be obtained by sampling hospital OR log books. This would probably be less expensive than surveying surgeons.

c. Problem Procedures

Conceptually, all procedures can be updated whether a "problem" or not. (Problem procedures are typically: new; represent rapidly changing technologies, infrequently performed). But again it is likely to be an expensive proposition to collect new data.

d. Differences by Specialty and Geographic Area

These can be readily developed with a data file sufficiently large to be disaggregated by specialty and geographic area.

e. Local Implementation

Individual Carriers could easily implement this system, but would have to make substantial investments in new data collection in order to obtain information on physicians' time inputs.

f. More Suited to Some Procedures?

An RVS based on physicians' time would be more suited to procedures which use either little or no non-physician inputs or procedures for which the non-physician inputs have a constant relationship to physician time. For example, a physician-time-based RVS would work relatively well for anesthesiology procedures where reimbursement excludes non-physician costs and costs are proportional to physician time.

g. Other Costs

Patient and other medical costs are not reflected in an RVS based on physician time. Patients' costs are not included nor are other medical care costs such as hospital charges.

Summary

a. Strengths

The principal strength of a time-based methodology is that its foundation, the time spent on a particular procedure, is a clear and well-understood unit of measure. The method of converting time into a relative value scale is also relatively straightforward, although determining the magnitude of the appropriate base value can be much more complex. The data collection is conceptually easy and, if enough data are collected, specialty- or region-specific RVSs can be computed. Because of problems with measuring nonphysicians costs, however, this method may be best suited for surgical procedures where physicians bear only a small proportion of the nonphysician costs.

b. Weaknesses

The greatest weakness of the method is that it measures only one input, physician time. Thus, there is great potential for distortion. Furthermore, a method of constructing "base values" which accurately reflect nonphysician costs and/or patient severity is difficult both to design and to implement. Data collection, while conceptually simple, would be expensive to carry out.

MICRO-COSTING TECHNIQUES

Description

Micro-costing is a general method of resource cost measurement whose purpose is to allocate the costs of a multi-product organization to the particular services or products of interest. The medical practice and hospital department have been treated as multi-product firms and have been analyzed by micro-costing techniques (Macdonald and Reuter, 1973; Kahn, Wirth, and Perkoff, 1978; Orkand, Jagger, and Hurwitz, 1977).

Micro-costing uses work-sampling and time-and-motion studies (Barnes, 1958) to obtain estimates of the relative burden on productive resources required by the different outputs of the practice or department. Work-load measurement techniques (College of American Pathologists, 1981), which focus on the input of personnel time to various services, constitute partial micro-costing, because non-personnel costs are not the measured in these studies and personnel resources are measured in time units, not dollar costs. In a full micro-costing study, the contribution of capital as well as labor inputs to the products of interest would be estimated and valued in commensurate units (i.e., dollars).

Micro-costing is useful when procedures vary widely in the amount required of direct non-physician resources. Equipment-intensive procedures, for example, would be allocated a larger proportion of equipment cost than would procedures using little equipment time. Laboratory and radiology services might well fall into this category. Even without equipment, micro-costing is useful when non-physician personnel can be readily substituted for physician time in the provision of services.

Micro-costing may also be useful in developing payment schedules based on case-mix or severity since the cost of some procedures is known to vary widely with the condition or disease state of the patient. Because it is based on randomly sampled observations of service delivery, the distribution of procedure costs across patients can be estimated.

To our knowledge, micro-costing techniques have not been used to establish relative values for medical procedures. However, the College of American Pathologists (CAP) has established a set of relative weights for clinical laboratory procedures which reflects time inputs of non-physician personnel (College of American Pathologists, 1981). These weights could be used as a basis for relative value construction. One would have to add estimates of both capital costs, including equipment and space, and physician costs to obtain complete micro-costing information.

Data

With the exception of the CAP workload measurement system, which is only a partial micro-costing approach, there are no generally available analyses of relative resource costs of medical procedures. New studies would have to be funded to develop such a method for other services. Because the productive inputs to the provision of services is likely to vary systematically with the setting of care, micro-costing studies would need to be applied in a variety of settings. Studies would also have to be repeated frequently to account of changing technology in the provision of services.

Micro-costing is an expensive technique. The authors of a 1971 study of micro-costing in a small number of hospitals estimated that a study for a single hospital department such as radiology would cost about $100,000. (Shuman, Wolf, and Perlman, 1973) As the desired level of generalizability of the results increases, so, too, do the costs of implementing the method. The findings of a micro-costing analysis of a few medical practices conducted in a few cities might not accurately represent the extent of interregional and cross-practice variation that actually exists in producing the services.

Technical Analysis Requirements

The micro-costing approach to RVS construction would require two kinds of technical skills: analytic skills in the design and implementation of studies; and judgement in the application of the estimates to a relative price system. Though much of the data collection could be conducted by trained clerical staff, the design, management, and analysis of such studies requires professional experience with work sampling methods in health care settings. Such skills generally reside in hospital industrial engineers. In addition, hospital accounting experts would be useful in assisting in the identification of input costs.

The results of such studies generally would not be directly applicable to a pricing system. Decisions about the appropriate levels of aggregation of cost data across patient groups and settings of care would be required. Committees consisting of knowledgeable practitioners and other kinds of health care professionals would be needed to offer this kind of judgement. Thus, to be used in a pricing system, the micro-costing approach would include the use of consensus development techniques. Individuals with some judgement and credibility would be essential if an acceptable pricing system were to be developed.

Tentative Characteristics of Resulting RVS

At its best, an RVS based on micro-costing techniques would reflect the relative value of resources required to produce a unit of different services. However, it would not reflect the relative contribution of each procedure to total patient costs. Some procedures, for example, require auxilliary services such as extra days of hospital care, drugs, and other ancillary services. These external costs do not enter the micro-costing framework.

Updating an RVS to account for new procedures or for new technological approaches to existing procedures would require new studies, with their attendant costs. Thus, the approach is not efficient in dealing with changing technology. However, micro-costing techniques can be easily adapted to handle problems associated with infrequently performed procedures. Precise estimates of the relative costs of providing these procedures can be obtained, though the existence of rare procedures implies that studies must be conducted over extended periods of time, thus further raising their costs.

Because micro-costing requires direct observation, the cooperation of medical practices and institutions is necessary. Such cooperation is often difficult for a third-party-payer to obtain. Consequently, these studies are probably best handled by professional societies.

Micro-costing is best suited for capital-intensive procedures and for procedures with a significant involvement of non-physician personnel. Among the possible classes of procedures are radiology, clinical laboratory and respiratory therapy.

Summary

Strengths

In summary, the micro-costing technique can provide an accurate measure of the relative cost of inputs to equipment-intensive procedures and procedures with heavy involvement of non-physician personnel. It can also be used to identify systematic differences in costs among patients and settings of care.

Weaknesses

The approach does not take account of resource costs external to the physician's practice. The major weakness of the approach is that it is extremely costly to implement and would require frequent updating and extensive on-site observation.

STATISTICAL COST FUNCTIONS

Description

One of the basic results of micro-economic theory is that resources are allocated efficiently when for all i products,

(1) $AR_i = MR_i = P_i = MC_i = AC_i$,

where P_i is the market price of the ith product, MC_i is its marginal cost, AC_i is its average cost, and AR_i and MR_i are average and marginal revenue, respectively. Economic theory also predicts that this will occur naturally in a world of perfect competition with complete information. In fact, it is precisely this notion which underlies the use of relative prices to construct relative values, since in competitive equilibrium prices accurately reflect both resource costs and consumers' valuations.

If, however, markets are not competitive, or information is imperfect, or insurance distorts private preferences and choices, then equation (1) is likely to be replaced by a set of inequalities,

(2) $AC_i = P_i > MC_i = MR_i < AR_i$.

In this case, no producer receives pure economic profits, since $AC_i = P_i$, but firms combine resources inefficiently, since average cost exceeds marginal cost. In addition, relative values based on relative prices would not reflect the relative costs of procedures under long-run efficiency.

Consequently, another approach to constructing relative values is to estimate statistical cost functions and compute the point at which the long-run efficiency condition, $MC_i = AC_i$, is satisfied. (This is equivalent to identifying the value of AC at its minimum.) Setting $P_i = MC_i$ would simultaneously set both relative values and absolute price levels. Thus, in theory at least, this approach has considerable merit.

Until recently, a major stumbling block in estimating cost functions for physicians' practices was the multiproduct nature of the physician firm. Physicians combine their own time, the time of their employees, various types of equipment, and floor space to produce many different services. Over the last few years, however, theoretical work on cost function specification by Caves and Christensen (1976) and Brown and Christensen (1979) has led to the development of a fairly general multiproduct cost function which permits the recovery of key production function parameters. The general long-run function may be written as

(3) $C = F(Q_1, \ldots, Q_n, W_1, \ldots, W_m)$,

where Q_i = practice outputs, $i = 1, \ldots, n$, and

W_j = practice input prices, $j = 1, \ldots, m$.

This function can be specified in translog form as

(4)
$$\ln C = \alpha_o + \sum_i \alpha_i \ln Q_i + \sum_i \beta_i \ln W_i + \tfrac{1}{2} \sum_i \sum_j \gamma_{ij} \ln W_i \ln W_j + \tfrac{1}{2} \sum_i \sum_j \zeta_{ij} \ln Q_i \ln Q_j + \sum_i \sum_j \phi_{ij} \ln W_i \ln Q_j + \varepsilon$$

The marginal cost of each output, shown by equation (5), can be readily computed from equation (4).

(5)
$$MC_i = \frac{\partial C}{\partial Q_i} = \frac{\hat{C}}{\hat{Q}_i} \cdot (\alpha_i + \sum_j \zeta_{ij} \ln Q_j + \sum_j \phi_{ji} \ln W_j)$$

The marginal cost of the ith output depends on the levels of the other outputs and all of the input prices. C/Q_i is the value of the average cost function at the point at which its slope with respect to Q_i is zero, i.e., the minimum point of the average cost function.

In principle, then, one could compute MC_i for each relevant output. Letting input prices vary by region and/or community size would enable one to compute different relative values for different types of areas. This is justified because the efficient combination of inputs will depend on relative input prices. One could also estimate specialty specific cost functions.

In practice, however, estimation of multiproduct statistical cost functions faces a number of potentially severe limits. The first is the sheer number of outputs a physician typically produces. Current procedure coding schemes include roughly five to six thousand separate procedures, each of which would presumably require a relative value. Incorporating this many outputs into a cost function is clearly not feasible. Thus, at a minimum, the number of outputs would have to be sharply compressed. Even if one concentrated on a highly aggregated output set (for example, various types of visits, which account for between 60 and 75 percent of revenues for most specialties, numbers of radiology, pathology, and other procedures, and number of operations), this would still leave more than a dozen output measures. The robustness and statistical reliability of multiproduct cost functions with this many output measures are not well known.

A second problem, also of a statistical nature, is that the output levels are chosen by the physician, i.e., the Q_i are endogenous. Unless some type of simultaneous equation estimation method is used, the coefficient values, which are critical to computing MC_i, may be biased. Unfortunately, the success of a simultaneous equation method depends to a large extent on the quality of the instruments used to replace the Q_i. If suitable instruments cannot be constructed, then the statistical reliability of the coefficients is likely to be low.

A third, and perhaps the most severe problem is that one of the most important data elements required to estimate cost functions is typically unobservable: the cost of the physician's time, which should be the major component of total costs. Although a number of surveys have information on physicians' net incomes it is clear that net income will typically include the financial returns from being an entrepreneur and owner of capital, as well as the cost of the physician's time input. Consequently, a computed hourly wage will not be a good measure of the physician's cost of own time. It might be possible to impute estimates based on data for physicians who are employees. However, employee-physicians are only a very small proportion of all physicians, and estimates based on their wages are likely to be poor proxies for physicians in general. Thus, errors of measurement for this key variable are likely to be another source of bias.

There are no relative value scales based on estimates of statistical cost functions. In fact, there do not appear to be any multiproduct cost functions estimated for physicians' practices. In general, cost function estimation for physicians' practices is a largely unmined area of research, to a large extent, one would guess, because of the inherent data problems mentioned above. Berry (1981) has reported single-product, trans-log cost functions for a number of specialties. Earlier single-output cost function estimates used total visits or patient encounters as the unit of output and very simple functional forms (Ernst and Schwartz, 1974; Newhouse, 1973; Yett, 1967). These studies were interested primarily in the question of economies of scale in physicians' practices.

Application of the multiproduct, trans-log cost function approach to health care providers is fairly recent. The General Accounting Office used this method successively in an analysis of the economic performance of HMOs (GAO, 1980). However, their study assumed that HMOs produce only three outputs, physicians' services, allied health professionals' services, and hospital services. There have also been some studies which estimated trans-log cost functions for hospitals, again using a small number of output measures (Roddy, 1980; Cowing and Holtman, 1980). All of these studies have produced plausible estimates of key parameters. Thus, the methodology appears to be feasible under fairly limited conditions. (These conditions are less restrictive, though, than those imposed upon cost function estimates for physicians' practices.)

Data Requirements

An ideal data base for estimating statistical cost functions should contain information on the quanitities of all of the outputs produced by the practice, the practice's total costs, including both implicit and explicit payments for physicians' time inputs, and the prices of all of the inputs employed, again including the price of a unit of physician input. Measuring the first component is complicated by the sheer number of different practice outputs. Obtaining the second and third parts of the ideal data base is difficult primarily because of the problems of measuring both the total and unit cost of physicians' own-time inputs. A secondary problem is measuring the costs and prices of capital inputs, typically office space and durable equipment.

No data base exists which meets all of these requirements. Of available data bases, two, HCFA's Practice Cost Survey and the AMA's Periodic Survey of Physicians, contain some of the required data elements. Both have reasonably good information on practice costs and some input prices. (The total and unit costs of physicians' time has to be imputed from data on physicians' net incomes, hours of work, and other characteristics.) Both also contain information on the quantities of a small number of practice outputs. The HCFA survey reports five types of office visits, the number of operations, and the proportion of visits with x-rays, lab tests, or injections. None of these outputs are identfied by a specific procedure code. The AMA survey, which varies somewhat from year to year also asks physicians to report quantities provided for a small number of visits and procedures performed. The 1980 survey, for example, asked for only three types of patient visits (office, hospital, and all other facilities). The 1978 survey asked for the numbers of three types of visits and eight other procedures, identifying procedures by their CRVS procedure code. In general, though, data bases which might have detailed data on practice outputs typically do not include information on costs or input prices. Conversely, data sets which do have cost and input price data usually contain only very limited output information.

The costs of building an ideal data base would be large, since direct survey of physicians' practices would be necessary. Expertise in obtaining cost and input price data for physicians' offices is fairly well developed. Developing a method of obtaining data on outputs would probably require a substantial design and development effort. Log diaries, examination of billing records, or physician (or staff) estimates of outputs provided are alternative methods which might be considered. A potential problem which would need to be explored is the probable lack of conformance between the time period for which the cost data apply and the time period over which the outputs were provided. All in all, it seems that a substantial research and development phase would be needed to explore the methodology in general, develop and test data collection strategies, and test the sensitivity of cost function parameter estimates to different types of data (numbers of observations and types of information collected).

Once the desired methodology were developed, the needed data could be obtained periodically by direct survey. The number of physicians surveyed, and hence the total cost of collecting the data, would depend on whether only a single RVS was to be developed, or several RVSs varying by specialty, region, community, size, and possibly other characteristics as well. For the purpose of providing a benchmark, HCFA's survey of physicians' costs covers about 10,000 physicians at a cost of approximately $700,000 per year.

Technical Analysis Requirements

Under current conditions, this approach requires a high level of economic, econometric, and survey research expertise. Both the data collection and statistical analysis phases require substantial research and development, ideally by experienced researchers. Once the method was tested and more fully developed, it could in principle be converted into routine data collection and analysis performed by less highly trained persons working with fixed survey instruments, coding manuals, and computer programs. However, higher level analysts would still be required to monitor ambiguous responses to the data collection and to adjust the methodology, if necessary, in response to changing conditions.

Tentative Characteristics

As noted in the descriptive section, an RVS constructed from statistical cost functions would, under ideal conditions, reflect both resource costs and patients' valuations. Even under the best conditions feasible, however, this is unlikely to occur because the costs incurred by a physician's practice understate total costs in two important ways. First, costs born by hospitals, clinics, nursing homes, and other facilities in which the physician practices are not reported by physicians and are not likely to be identifiable without incurring unacceptable data collection costs. Second, costs born by patients, primarily the value of travel and waiting time, are also not reported and difficult to measure at best. The first problem could be dealt with by limiting the method only to procedures performed in the physician's office. (One would then have to identify the amount of time the physician spends in the office and divide the physician's earnings into office and nonoffice components. The latter would seem to be particularly difficult, especially

What is the importance of non-MD costs for reimbursement?

for surgeons and other specialists who derive a substantial portion of their incomes from procedures performed outside of the office.) The second problem, accounting for patients' costs, does not seem to have any ready solution short of a much more expensive data collection effort. As a result of these problems, a RVS constructed from statistical cost function estimates is not likely to be a good reflection of either the full social costs of providing different procedures or of patients' valuations of those procedures.

This approach also suffers from a number of other implementation problems, some of which were alluded to above. Each update of the RVS would require an expensive primary data collection effort. Furthermore, this cost would be directly related to the number of RVSs which were in use, i.e., specialty-, region-, and/or community-size-specific RVSs. The method would be extremely difficult to implement by individual carriers because of the data collection costs and substantial technical analysis requirements. This method is not well suited for constructing RVSs for procedures infrequently performed in physicians' offices. (Procedure-specific cost functions, and hence relative values, might be feasible to estimate if the procedure rather than the physician's office were the unit of analysis. This would be a highly speculative undertaking, however.) Finally, the method would be difficult to understand, if not incomprehensible to almost all physicians, legislators, and insurors who might have to evaluate the RVS or be affected by it. This would severely hamper the use of the RVS as a benchmark in any type of provider and insurer/government negotiations or bargaining over relative fees.

<u>Summary</u>

The primary advantage of a RVS based on statistical cost functions is its relationship to prices which would prevail under conditions of a competitive equilibrium. This strength would appear to be outweighed by many weaknesses in the actual development and implementation of the method. Available data would permit at best only a testing of the statistical methods. There may be some value, however, in estimating specialty- or area-specific cost functions for a small number of outputs. Estimates of minimum points on average cost functions could then be used to address the question of whether there should be one or multiple RVSs. For example, if $MC_j = AC_j$ = some constant value, K, over j regions, then this would strengthen the argument for using only a single national RVS rather than separate regional RVSs. Building a data file for more intensive development would be a very costly undertaking, as would be collecting data for subsequent updates.

CONSENSUS DEVELOPMENT

Description

Consensus development refers to a continuum of techniques whose purpose is to arrive at relative values through pooling of expert judgements. Techniques range from ordinary committees of experts who follow informal or formal rules of order to specially structured face-to-face group processes and even to anonymous pooling of opinions by delphi techniques (Hiltz and Turoff, 1978).

In a sense, all relative value construction methods must use consensus development at some point. To be adopted by an organization for example, a Board of Directors must vote on a proposed RVS. The California Relative Value Studies used committees as an adjunct to analysis of charge data and ultimately as a method for introducing new procedures.

The Delphi method (Dalkey, 1968) is just one of several approaches available to arrive at a group RVS. In contrast to simple questionnaires administered to a group of respondents, Delphi involves three to four successive rounds of anonymous questionnaires with feed-back of information to respondents between rounds. In the first round, for example, individuals might be asked to rate each procedure on a Lickert-type scale. The median rating would then be returned to the respondents with a request to perform a second rating. More elaborate Delphi questionnaires may ask respondents to list the considerations underlying their rating, and these qualitative judgements may be fed back to all respondents in successive rounds.

Delphi methods were originally developed for technological forecasting; they remain the method of last resort when empirically derived estimates are not possible. Experiments with the Delphi technique have shown that convergence to a "consensus" is common over three to four rounds of questionnaires, where consensus is defined as an acceptably low level of variation in estimates around the mean or median (Dalkey, 1968).

Whether Delphi performs better than face-to-face groups is subject to conflicting evidence. Campbell (1968) found Delphi groups to be more accurate at short-term economic forecasts than face-to-face encounters, but Farquhar (1970) found significantly better results in face-to-face groups required to solve a complex estimation task. However, neither of these studies compared the impact of group size on the accuracy or quality of the outcome of either kind of group. Because Delphi allows for participation of more experts than do face-to-face groups, which have a practical limit of about 15 (Filley, 1970), it may in reality be a more accurate technique.

Data

Consensus development approaches are flexible with respect to data needs because it is assumed that some information already resides in the experts themselves. However, supporting empirical data can be readily absorbed into the decision-making process. Thus, consensus development can accept virtually any level of information.

Technical Analysis Requirements

The requirements for technical analysis depend on the consensus development technique selected and the kinds of supporting empirical data that would be produced for the panels. More structured processes such as nominal groups, delphi, and computer conferencing require design by those experienced with these methods, whereas traditional committees require only modest staff support.

Characteristics of Resulting RVS

An RVS developed by consensus techniques can represent any concept desired by the sponsors, although the technique is most likely to be applied when a concept is difficult to measure as in the case of social preferences or resource costs. The designers of the system must know and be able to communicate to panel members the conceptual basis for RVS construction. Vague instructions are likely to induce unstable and unconfident responses (Scheibe, Skutsch, and Schofer, 1975).

The consensus approach is flexible and relatively inexpensive to implement. The RVS resulting from such an approach can be easily updated on a periodic basis, and can be modified as desired by regional or specialty-specific panels. However, there is a real danger of "social abuse" of the process due to the uncomfortable link between expertise in the technical issues and financial stake in the outcome of the process (Sackman, 1974). Even absent this conflict, panel composition is likely to affect the outcome. Thus, while the design of a consensus development process can admit many possibilities, each will have different implications for the resulting RVS.

The use of this family of techniques is most appropriate for cases in which an adequate data base does not exist to develop empirically derived estimates. In particular, consensus techniques would be valuable in establishing relative values for new procedures or for procedures with rapidly changing technologies; for subsets of visits or procedures with varying levels of complexity that cannot be approximated by physician times; and for conversion factors across broad procedure categories.

Summary

Consensus development techniques are flexible and relatively inexpensive, admitting any level of data analysis, from reliance on knowledge already held by experts to the production of supporting data. It also can be used for any concept of relative values, although the underlying concept should be explicit at the time of implementation.

Its very flexibility also implies potential instability and dependence on the composition of the expert panel (Dalkey, 1969; Bedford, 1972). It should be viewed as a method of last resort when empirically derived scales are not obtainable due to prohibitive costs or methodological obstacles. It is, therefore, most appropriate as a selective tool for establishing relative values for new procedures and for conversion rates across broad categories of procedures.

SOCIAL PREFERENCE APPROACH

Description

The social preference approach views the RVS (or a fee schedule based on such a scale) as an instrument for allocating health care resources in a socially desirable manner. The identification of "socially desirable" patterns of service provision is not straightforward, but as a guiding principle, one might expect a desirable allocation of health services to equalize the ratio of net social benefit to social cost across all procedures.* A relative value scale that encourages providers to offer this socially efficient mix of services is the goal of the approach.

Though numerous examples exist of procedures that are allegedly over- and under-used relative to the socially efficient level, the extent to which alterations in relative fees can infuence these disparities is not well understood. Moreover, there is some disagreement on how the social benefits and costs of medical procedures should be measured, or indeed whether they can be measured at all.

Several public and third party payers have attempted to use physician fee policies to influence the allocation of resources in directions considered to be more socially efficient. Denial of payment for procedures performed in unnecessarily costly settings is one example. Per-case reimbursement is another. But the manipulation of the structure of fees for the expressed purpose of encouraging more socially efficient patterns of care has occurred infrequently. Perhaps the best known effort is the California Medicaid program's 1976 adoption of a fee schedule that radically altered the relative rate of payment to physicians between office visits and other billable procedures in favor of office visits. An evaluation of the impact of this policy instrument on the use of services is currently underway.

---

*To the extent that social benefits include the distributional impacts of any allocation, which requires the interpersonal comparison of individual utilities, the social preference approach is also concerned with these issues.

Two other examples come from abroad. The provincial health plan of Quebec, Canada sets relative values so that rates of return to physician training are roughly equalized across specialties. In effect, the Province's social preference is to preserve the existing mix of specialists. The government of West Germany focuses on rates of increase in expenditure. If expenditures for physicians' services grow faster than planned, then RVSs (or conversion factors) are scaled back in order to meet expenditure targets. This type of manipulation could be performed for selected types of procedures, e.g., surgeries, or across-the-board for all procedures.

Data

Data needs depend, first, upon the amount of evidence required to support construction of a relative value scale based on social preferences, and, second, on the way in which the approach is to be applied to scale construction. If expert opinion is considered adequate for establishing relative social values among procedures, then committees or panels can be convened with relatively low costs. At the other extreme, if determination of relative values must be made on the basis of evidence about effectiveness and costs that meets high standards of scientific validity, data requirements could well be prohibitive. The level of detail of application is also important. If one were to establish relative values across broad classes of procedures or perhaps across a limited set of individual procedures, the need for supporting evidence might be kept at a reasonable level. Application of a social preference approach across all individual procedures would be impractical in the extreme.

Data should be available to monitor the impacts of a change in relative values on the use of services. Because the process of establishing relative values is likely to be imprecise, feedback on the effects of changes could be used to adjust values over time. The acceptability of resulting rates of use of procedures can be ascertained if such data are available. Third party claims files would constitute the basic elements of required data.

Technical Analysis Requirements

These requirements go hand in hand with data needs. If panels or committees are to be convened, those with skills in the technology of consensus development would be needed. If evidence on the benefits, risks, and costs of alternative procedures is to be generated, then sophisticated skills in clinical and economic evaluations of medical procedures are required.

Characteristics of a Social Preference - Based RVS

An RVS based on social preferences would represent the best estimate of, or opinions about, the relative net social benefits of medical services. The concept captures not only the direct and indirect costs of performing procedure but also the value to society of the services provided.*

The ease with which the method can be implemented depends upon the standards of evidence required for construction of relative values. Using a consensus development approach, the method could be implemented at minimal cost and is amenable to refinement over time or by geographic region or specialty. Most third party payers either have the capacity to organize such efforts or have access to organizations that can implement the system.

The technique can be applied relatively easily to the establishment of the RVS conversion rates among broad classes of procedures or for services delivered in different settings. For example, if one wanted to encourage

---

*Social value would include external benefits as well as the benefits accruing to the recipient directly.

ambulatory surgery as a substitute for inpatient surgery, a premium (higher relative value) could be established for any such procedure performed in outpatient setting. The method may also be useful in isolated instances in establishing the relative value of two or more procedures that are good substitutes for one another.

The method would not be particularly useful for new procedures or for those with fast changing technologies for two reasons. First, the social benefits and costs of new procedures are generally not well documented nor understood when a procedure is first introduced. Second, the early evidence that does exist is generally based on the experience of a few practitioners and institutions. A relative value system should reflect the experience with a procedure when it reaches the point of standard practice.

Summary

The social preference approach is based on a view of the fee schedule as an instrument of resource allocation. It is potentially easy to implement, especially in establishing conversion rates across broad categories of services or across services rendered in different settings. It is, however, an imprecise tool in that the determination of social benefits and social costs of procedures is conceptually difficult and may be prohibitively expensive. When coupled with a system for monitoring the impact of fee schedules based on relative social values or the relative rates of use of services, the approach can be a valuable health policy tool.

The approach is limited by the lack of knowledge about social benefits and costs in many critical areas. For example, although some have advocated that the structure of physicians' fees be "twisted" in the direction of greater payment for provision of "cognitive services" relative to technical procedures, evidence regarding the patient benefits to be derived from a reallocation of resources in this direction does not exist.

III. RECOMMENDATIONS

This section recommends which methods should receive further consideration in developing the final research plan for the project. Our primary goal at this point is to eliminate or deemphasize those methods which appear to be least promising on either conceptual, technical, or data grounds. It is important to make such decisions at this time in order to avoid diluting the resources available for analysis of the more promising methods. Our overall recommendation is that charge-based methods, physicians' time methods, and a combination of delphi and social preference function methods should receive more detailed consideration. Various micro-costing and statistical cost function approaches have considerable theoretical appeal, but also face the most serious technical and data requirements problems. We do not feel that an RVS could be developed using either of these two approaches within the resource constraints of this project. At best, these methods should receive only limited exploratory/developmental consideration.

Charge Methods

Charge data collected as part of the current claims payment and fee screen development processes are a promising source for developing RVSs. The marginal costs of acquiring such data would be low. Computer processing costs may be high, but would probably not be substantially different from the current costs of developing carrier-specific fee screens. Charge data offer substantial detail and considerable flexibility in terms of tailoring RVSs to specific situations, e.g., specialty, region, and/or community size. The primary conceptual drawback of a charge-based method is that charges may contain distortions induced by the uneven pattern of insurance coverage across areas, specialties, and different types of procedures. These problems may be amenable to adjustment in conjunction with the delphi-social preference function approaches.

We propose focusing on three different charge data files: the Urban Institute's California Medicare and Medicaid Claims files, Medicare's prevailing charges (50th and 75th percentiles) by locality for 100 Medicare procedures, and the Health Insurance Association of America's survey of commercial insurors' claims for surgical procedures. In our Final Research Plan we will spell out the details of constructing and comparing alternative RVSs using these data. We will limit the anaysis to the 100 procedures for which Medicare prevailing charges are available plus a small number of "problem" procedures. We will explore differences in RVSs based on alternative points on the distribution of charges, region, specialty, community size, claims source, and procedure type.

Cost Methods

Of the three cost methods examined, we feel that the greatest emphasis should be given to constructing an RVS using physicians' time or time plus a base value. This method may be criticized on two grounds: it focuses on only one input to the production of physicians' services, and it does not account for patients' valuations. Nevertheless, there may be a number of procedures for which physicians' time is either the major input or other inputs are generally combined in a fixed relationship with physicians' time. Furthermore, time is an easily understood basis for building an RVS and it is easier to define and measure than the variables required for micro-costing or statistical cost functions.

As a general approach to constructing an RVS for a large number of procedures, both micro-costing and statistical cost functions appear to be both too expensive (in terms of data collection and data analysis requirements) and too shaky on conceptual grounds. However, each may have some utility in specific situations. In particular, micro-costing may be feasible for well-defined intermediate outputs such as lab tests and some radiology procedures.

Statistical cost functions may be useful in comparing office-visit RVSs across specialties, regions, and community sizes. The question one would ask is, essentially, whether the minimum points of the average cost functions for different types of office visits vary among specialties, regions, or community sizes. If the underlying production and cost functions are similar (except for scalar transformations), then this would strengthen the case for a single RVS.

We recommend concentrating primarily on the analysis of the USC-Mendenhall data and the development of time-based RVSs for a selected number of procedures. Although detailed procedure coding is not available, this file still has potential value because it measures time spent with patients in different settings and under varying complexity-urgency combinations. Thus, it should be possible to construct relative values for visits of different complexity by setting (office, inpatient hospital, and outpatient) and for different specialties, regions, and community sizes. Since visit codes account for the bulk of expenditures for physicians' services, even this limited RVS could be highly pertinent to many policy decisions. Furthermore, comparing time-based with charge-based relative values for office visits may produce important information for evaluating each approach. In particular, if they are similar, then arguments for a charge-based system would be reinforced.

A second data base which we could examine is the National Ambulatory Medical Care Survey. Although not as rich as the USC-Mendenhall data, it does have some overlap in terms of office visit time by diagnosis and severity. It also has the advantage of being an ongoing, annual data collection effort. We recommend that it be used to validate to the extent possible RVSs constructed from the USC-Mendenhall data, and to examine the stability over time of relative time per visit.

We also recommend that analysis of micro-costing and statistical cost function methods be limited to no more than the following: a more detailed secondary evaluation of micro-costing methods than presented above; and exploratory estimation of multiproduct statistical cost functions using data from HCFA's Survey of Physicians' Practice Costs.

Consensus Methods

Consensus methods, both delphi and social preferences, do not appear to be well suited for constructing RVSs de novo. Neither method is equipped to process and evaluate large amounts of data for large numbers of procedures. Rather, these methods should be thought of as complements to one or more of the data-oriented methods described above. Their advantage is that they provide a distinctly human perspective--the ability to make broad imprecise comparisons and to apply values--than the more mechanistic approaches described above. We believe that this capability should be an important part of developing RVSs, but is probably incapable of doing so in and of itself. Therefore, the consensus methods should be evauated in conjunction with both the time- and charge-based methods.

Another distinguishing characteristic of the consensus methods is that their application is intrinsically tied to the goal of the process. Selecting panel members, gathering data to be examined, and collecting documents for review all depend on the specific goal or objective sought. Thus, in order to analyze further consensus methods' feasibility and utility in constructing RVSs, it is necessary to make choices about the questions to be addressed.

We recommend three areas which could be usefully addressed by consensus methods. The first is identifying and specifying social preferences. The second is assessing relative values across different classes of procedures. Should medical, surgical, pathology, and radiology procedures all be measured along a single RVS? If yes, what should the base be and how can the conversions be made? The third is evaluating differences among different RVSs or relative values within procedure classes for specific RVSs. In both the second and third uses, one or more RVSs developed by some other means are inputs to the delphi process.

REFERENCES

1. Barnes, Ralph M. Motion and Time Study. New York: John Wiley & Sons, Inc., 1958.

2. Bedford, Michael T. "The Value of Competing Panels of Experts and the Impact of 'Drop-outs' on Delphi Results," Delphi: The Bell Canada Experience, Bell Canada, October 1972.

3. Berry, Robert. "A Review of the Medicare Economic Index: An Analysis of Variations in Practice Costs Among Specialties," in Greenspan, Nancy, ed., Issues in Physician Reimbursement, USDHHS, HCFA Pub. No. 03121, August 1981.

4. Brown, R., Caves, D., and Christensen, L. "Modelling the Structure of Cost and Production for Multi-Product Firms." Southern Economic Journal, July 1979, pp. 256-267.

5. Campbell, R.M. "The Delphi Technique: Implementation in the Corporate Environment," Management Services 5 (November-December 1968):37-42.

6. Caves, D.W. and Christensen, L.R. "Modelling the Structure of Production in the U.S. Railroad Industry." Paper presented at the winter meetings of the Econometric Society, September 1976.

7. College of American Pathologists. "Laboratory Workload Recording Method," 1981.

8. Cowing, T.G. and Holtmann, A.G. "The Estimation of Hospital Technologies Using a Short-Run Multi-Product Cost Function." Working Paper 50-1, Department of Economics at Binghamton, June 1980.

9. Dalkey, N.C. Experiments in Group Prediction. The Rand Corporation, P-3820, March 1968.

10. Dalkey, N.C. The Delphi Method: An Experimental Study of Group Opinion. The Rand Corporation, RM-5888-PR, June 1969.

11. Ernst, R.L. and Schwartz, H.J. "Determinants of the Costs of Physicians' Services," USC, Human Resources Research Center, January 1974.

12. Farquhar, J.A. A Preliminary Inquiry into the Software Estimation Process. The Rand Corporation, RM-6271-PR, August 1970.

13. Filley, A.C. "Committee Management: Guidelines from Social Science Research," California Management Review 13(1) Fall 1970:13-21.

14. General Accounting Office. "Health Maintenance Organizations Can Help Control Health Care Costs," PAD-80-17, May 6, 1980.

15. Kahn, L., Wirth, P., and Perkoff, G. "The Cost of a Primary Care Teaching Program in a Prepaid Group Practice," Medical Care 16 (January 1978).

16. Macdonald, L. and Reuter, L.F. "A Patient-Specific Approach to Hospital Cost Accounting," Health Services Research (Summer 1973).

17. Newhouse, J. "The Economics of Group Practice," Journal of Human Resources, VIII Winter 1973:35-56.

18. Orkand, D., Jaggar, F.M. and Hurwitz, E. "Comparative Evaluation of Costs, Quality, and System Effects of Ambulatory Surgery Performed in Alternative Settings," Final Reporet for HCFA, NTIS HRP-0025749, December 1977.

19. Roddy, D.J. "Subsidization and Inpatient/Outpatient Rate Structure in the Hospital Industry." Unpublished paper, July 1980.

20. Sackman, H. Delphi Assessment: Expert Opinion, Forecasting, and Group Process. The Rand Corporation, R-1283-PR, April 1974.

21. Sheibe, M., Skutsch, M. and Schofer, J. "Experiments in Delphi Methodology," The Delphi Method: Techniques and Applications. Reading, Mass.: Addison-Wesley Publishing Company, 1975.

22. Shuman, L., Wolfe, H. and Perlman, M. "Model for Hospital Microcosting," Industrial Engineering July 1973:39-43.

23. Yett, D. "An Evaluation of Alternative Methods of Estimating Physicians' Expenses Relative to Output." Inquiry, IV, March 1967:3-27.

CMS LIBRARY
3 8095 00013778 2